1

SEQUENCE LISTING

(1) GENERAL INFORMATION: (i) APPLICANT: (A) NAME: Studiengesellschaft Kohle mbH (B) STREET: Kaiser-Wilhelm-Platz 1 (C) CITY: Muelheim an der Ruhr (E) COUNTRY: Germany (F) POSTAL CODE (ZIP): 45470 (ii) TITLE OF INVENTION: A Process for the Preperation and Identifiation of Novel Hydrolases Having Improved Properties (iii) NUMBER OF SEQUENCES: 21 (iv) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: Paten | Release #1.0, Version #1.30 (EPO) (2) INFORMATION FOR SEQ ID NO: 1 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 30 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: unknown (D) TOPOLOGY: linear (ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "synthetic DNA" (xi) SEQUENCE DESCRIPTION: SEO ID NO: 1: GCGCAATTAA CCCTCACTAA AGGGAACAAA 30 (2) INFORMATION FOR SEO ID NO: 2: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: unknown (D) TOPOLOGY: linear (ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "synthetic DNA" (xi) SEQUENCE DESCRIPTION: SEO ID NO: 2: GCGTAATACG ACTCACTATA GGGCGAA 27 (2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1049 base pairs(B) TYPE: nucleic acid

(C) STRANGEDNESS: unknown (D) TOPOLOGY: unknown (ii) MOLECULE TYPE; DNA (genomic) (ix) FEATURE: (A) NAME/KEY:\CDS (B) LOCATION: 85..1017 (ix) FEATURE: (A) NAME/KEY: hat peptide (B) LOCATION: 163..1017 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3: 60 GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG TAT CTG CTC CCC CTC 111 Met Lys Lys Lys Tyr Leu Leu Pro Leu **-**26 -25 GGC CTG GCC ATC GGT CTC GCC GCT CTC GCT GCC AGC CCT CTG ATC CAG Gly Leu Ala Ile Gly Leu Ala Ala Ser Pro Leu Ile Gln 159 GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC 207 Ala Ser Thr Tyr Thr Gln Thr Lys \Tyr Pro Ile Val Leu Ala His Gly ATG CTC GGC TTC GAC AAC ATC CTC GGG GTC GAC TAC TGG TTC GGC ATT Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile 20 25 CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GTC Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Val 303 35 AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG 351 Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln 399 GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile GGC CAC AGC CAC GGC GGG CCG ACC ATC CGC TAC GTC GCC GCC GTA CGT Gly His Ser His Gly Gly Pro Thr Ile Arg Tyt Val Ala Ala Val Arg 447 CCC GAC CTG ATC GCT TCC GCC ATC AGC GTC GGC GCC CCG CAC AAG GGT 495 Pro Asp Leu Ile Ala Ser Ala Ile Ser Val Gly Ala Pro His Lys Gly 105 TCG GAC ACC GCC GAC TTC CTG CGC CAG ATC CCA COG GGT TCG GCC GGC 543 Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly GAG GCA GTC CTC TCC GGG CTG GTC AAC AGC CTC GGC GCG CTG ATC AGC 591 Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser

	CTT Leu 145															639
	TCG Ser															687
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr 180	TCG Ser	ALL a	TGC Cys	GGC Gly	GAA Glu 185	GGC Gly	GCC Ala	TAC Tyr	AAG Lys	GTC Val 190	AAC Asn	735
	GTG Val															783
	GAT Asp															831
AAC Asn	GGC Gly 225	ACC Thr	GCC Ala	AAC Asn	GAC Asp	GGC Gly 230	C. Leu	OTC Val	GGC Gly	ACC Thr	TGC Cys 235	AGT Ser	TCG Ser	CAC His	CTG Leu	879
	ATG Met															927
	CAG Gln															975
	TAC Tyr															1017
TAG	GACC	CCG	GCCG	GGGC	CT C	GGCC	CGGG	c cc	\							1049
(2)	TND	00117	m T AN	EOD	250	TD	NO.	۸.	,	\						
(2)	INF									\						
		(.	A) L B) T	ENCE ENGT YPE: OPOL	H: 3 ami	11 a no a	mino cid									
				LE T CE D					ID N	0: 4						
	Lys -25		Lys	Tyr	Leu	Leu -20	Pro	Leu	Gly	Leu	Ala -15	Ile	Gly	Leu	Ala	
Ser -10	Leu	Ala	Ala	Ser	Pro		Ile	Gln	Ala	Ser 1	Th	Tyr	Thr	Gln 5	Thr	
Lys	Tyr	Pro	Ile 10	Val	Leu	Ala	His	Gly 15	Met	Leu	Gly	he	Asp 20		Ile	
Lev	Gly	Val 25		Tyr	Trp	Phe	Gly 30		Pro	Ser	Ala	Lau 36	Arg	Arg	Asp	

Gly Ala Gln Val Tyr Val Thr Glu Val Ser Gln Leu Asp Thr Ser Glu 45 Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala Ile Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser\Ala Gly Glu Ala Val Leu Ser Gly Leu 125 Val Asn Ser Leu Gly Ala Leu Ale Ser Phe Leu Ser Ser Gly Gly Thr Gly Thr Gln Asn Ser Leu Gly Set Leu Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Atn Gly Val Ser Tyr Tyr Ser Trp 190 Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe 205 Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly 215 225 Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn 245 Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser Val tyr Arg Gln His Ala Asn

(2) INFORMATION FOR SEQ ID NO: 5:

Arg Leu Lys Asn Ala Ser Leu

280

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1049 base pairs

285

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 85..1017

(ix) FEATURE:

						mat_ .63										
	(xi)	SEÇ	UENC	E DE	GRI	PTIC	N: S	SEQ 1	D NO): 5:						
GGAI	cccc	CG G	TTCI	ccce	G ∤ A≱	AGGAI	TCGC	GCC	SATGO	CTG	GCAG	GACG	CG C	CCCI	CGGCC	60
CCAT	CAAC	CT G	AGAT	'GAGA	AA d	M		Lys I				eu I		ccc c Pro I		111
														ATC Ile		159
														CAC His		207
ATG Met	CTC Leu	GGC Gly	TTC Phe	GAC Asp 20	AAC Asn	ATC	CTC eu	GGG Gly	GTC Val 25	GAC Asp	TAC Tyr	TGG Trp	TTC Phe	GGC Gly 30	ATT Ile	255
														GAA Glu		303
AGC Ser	CAG Gln	TTG Leu 50	GAC Asp	ACC Thr	TCG Ser	GAA Glu	GTC Val 55	CGC Arg	GGC Gly	GAG Glu	CAG Gln	TTG Leu 60	CTG Leu	CAA Gln	CAG Gln	351
														CTG Leu		399
														GTA Val		447
														AAG Lys 110		495
														GCC Ala		543
GAG Glu	GCA Ala	GTC Val 130	CTC Leu	TCC Ser	GGG Gly	CTG Leu	GTC Val 135	AAC Asn	AGC Ser	TC Leu	GGC Gly	GCG Ala 140	CTG Leu	ATC Ile	AGC Ser	591
														TCG Ser		639
GAG Glu 160	TCG Ser	CTG Leu	AAC Asn	AGC Ser	GAG Glu 165	Gly	GCC Ala	GCG Ala	CGC Arg	TTC Phe 170	A)sn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro 175	687

CAG GGC ATC CCC ACC TCG GCC TGC GGC GAA GGC GCC TAC AAG GTC AAC Gln Gly Ile Pro Thr ter Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn 185	735
GGC GTG AGC TAT TAC TOC TGG AGC GGT TCC TCG CCG CTG ACC AAC TTC Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe	783
CTC GAT CCG AGC GAC GCC TTC CTC GGC GCC TCG TCG CTG ACC TTC AAG Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys 210 220	831
AAC GGC ACC GCC AAC GAC GAC GTG GTC GGC ACC TGG AGT TCG CAC CTG Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu 225 235	879
GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val 240 250 255	927
AAC CAG GTC TTC GGC CTC ACC ACC CTG TTC GAG ACC AGC CCG GTC AGC ASN Gln Val Phe Gly Leu Thr Sen Leu Phe Glu Thr Ser Pro Val Ser 265 270	975
GTC TAC CGC CAG CAC GCC AAC CGC OTG AAG AAC GCC AGC CTG Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu 275 280 285	1017
TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC	1049
(2) INFORMATION FOR SEQ ID NO: 6:	
(2) INFORMATION FOR SEQ ID NO: 6: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: Met Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3ll amino acids (B) TYPE: amino acids (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -26 -25 Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -26 -25 Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr -10 -5 Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -15 Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr -10 Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile 10 Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acids (C) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6: Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -26 -25 Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr -10 Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile 10 Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp 35 Gly Ala Gln Val Tyr Val Thr Glu Val Ser Gln Leu Asp Thr Ser Glu	

Thr Ile Arg Tyr Val A \oint_{Φ} Ala Val Arg Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr Gly Thr Gln Asn Ser Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys VA Asn Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe 205 Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn 245 Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn 270 Arg Leu Lys Asn Ala Ser Leu 280

- (2) INFORMATION FOR SEC ID NO: 7:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1049 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: unknown
 - (D) TOPOLOGY: unknown
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 85..1017
 - (ix) FEATURE:
 - (A) NAME/KEY: mat_peptide
 - (B) LOCATION: 163...1017
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCA?	CAA	CCT (GAGA!	rgagi	A CA	1	ATG A Met 1	Lys 1	AAG A	AAG :	TCT (Ser :	CTG (Leu 1	CTC (Leu I	CCC (Pro 1	CTC Leu	111
												CCT Pro -5				159
GCC Ala	AGC Ser 1	ACC Thr	TAC Tyr	ACC Thr	CAG Gln 5	ACC Thr	AAA Lys	TAC Tyr	CCC Pro	ATC Ile 10	GTG Val	CTG Leu	GCC Ala	CAC His	GGC Gly 15	207
ATG Met	CTC Leu	GGC Gly	TTC Phe	GAC Asp 20	AAC Asn	ATd Ile	CTT Leu	GGG Gly	GTC Val 25	GAC Asp	TAC Tyr	TGG Trp	TTC Phe	GGC Gly 30	ATT Ile	255
												GTC Val				303
												TTG Leu 60				351
GTG Val	GAG Glu 65	GAA Glu	ATC Ile	GTC Val	GCC Ala	CTC Leu 70	AGC Ser	dry gcc	CAG Gln	CCC Pro	AAG Lys 75	GTC Val	AAC Asn	CTG Leu	ATC Ile	399
												GCC Ala				447
CCC Pro	GAC Asp	CTG Leu	ATC Ile	GCT Ala 100	TCC Ser	GCC Ala	ACC Thr	AGC Ser	GTC Val 105	GGC Gly	GCC Ala	CCG Pro	CAC His	AAG Lys 110	GGT Gly	495
TCG Ser	GAC Asp	ACC Thr	GCC Ala 115	GAC Asp	TTC Phe	CTG Leu	CGC Arg	CAG Gln 120	ATC Il	CCA Pro	CCG Pro	GGT Gly	TCG Ser 125	GCC Ala	GGC Gly	543
GAG Glu	GCA Ala	GTC Val 130	CTC Leu	TCC Ser	GGG Gly	CTG Leu	GTC Val 135	AAC Asn	AGC Ser	CTC Leu	GGC Gly	GCG Ala 140	CTG Leu	ATC Ile	AGC Ser	591
												CTG Leu				639
												GCC Ala				687
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr 180	TCG Ser	GCC Ala	TGC Cys	GGC Gly	GAA Glu 185	GGC Gly	CC la	TAC Tyr	AAG Lys	GTC Val 190	AAC Asn	735
												CTG Leu				783

CTC Leu	GAT Asp	CCG Pro 210	AGC Ser	GAC Asp	GCC Ala	TTC Phe	CTC Leu 215	GGC Gly	GCC Ala	TCG Ser	TCG Ser	CTG Leu 220	ACC Thr	TTC Phe	AAG Lys	83
AAC Asn	GGC Gly 225	ACC Thr	GCC Ala	AAC Asn	GAC	GGC Gly 230	CTG Leu	GTC Val	GGC Gly	ACC Thr	TGC Cys 235	AGT Ser	TCG Ser	CAC His	CTG Leu	87
		GTG Val														92
		GTC Val														97
GTC Val	TAC Tyr	CGC Arg	CAG Gln 275	CAC His	GCC Ala	AAC Asn	CGC Arg	CTG Leu 280	AAG Lys	AAC Asn	GCC Ala	AGC Ser	CTG Leu 285			101
TAGO	SACC	CCG (GCCGC	GGGC	CT CC	GC	Jeggg	c cc								104
(2)	INF	ORMA:	rion	FOR	SEQ	ID N	1 g ∶ 8	3:								
		(1	A) LI 3) T	ENGTI YPE:	CHAP H: 31 amir OGY:	ll ar	nin cid									
) MOI							ID N	0: 8	:					
	Lys -25	Lys	Lys	Ser	Leu	Leu -20	Pro	refu	Gly	Leu	Ala -15	Ile	Gly	Leu	Ala	
Ser -10	Leu	Ala	Ala	Ser	Pro ~5	Leu	Ile	Gln	Ala	Ser 1	Thr	Tyr	Thr	Gln 5	Thr	
Lys	Tyr	Pro	Ile 10	Val	Leu	Ala	His	Gly 15	Met	Leu	Gly	Phe	Asp 20	Asn	Ile	
Leu	Gly	Val 25	Asp	Tyr	Trp	Phe	Gly 30	Ile	Pro	Ser	Ala	Leu 35	Arg	Arg	Asp	
Gly	Ala 40	Gln	Val	Tyr	Val	Thr 45	Glu	Val	Ser	Gln	Leu 50	Asp	Thr	Ser	Glu	
Val 55	Arg	Gly	Glu	Gln	Leu 60	Leu	Gln	Gln	Val	d1 u	Glu	Ile	Val	Ala	Leu 70	
Ser	Gly	Gln	Pro	Lys 75	Val	Asn	Leu	Ile	Gly 80	His	Ser	His	Gly	Gly 85	Pro	
Thr	Ile	Arg	Tyr 90		Ala	Ala	Val	Arg 95		Asp	Leu	Ile	Ala 100		Ala	
Thr	Ser	Val 105		Ala	Pro	His	Lys 110		Ser	Asp	hr	A1a		Phe	Leu	
Arg	Gln 120		Pro	Pro	Gly	Ser 125		Gly	Glu	ı Ala	va) 130	Leu	Ser	Gly	Leu	

Val 135	Asn	Ser	Leu	e f	Ala 140	Leu	Ile	Ser	Phe	Leu 145	Ser	Ser	Gly	Gly	Thr 150	
Gly	Thr	Gln	Asn	Leu 155	Leu	Gly	Ser	Leu	Glu 160	Ser	Leu	Asn	Ser	Glu 165	Gly	
Ala	Ala	Arg	Phe 170	Asn	Ala	Lys	Tyr	Pro 175	Gln	Gly	Ile	Pro	Thr 180	Ser	Ala	
Cys	Gly	Glu 185	Gly	Ala	Tyn	Lys	Val 190	Asn	Gly	Val	Ser	Tyr 195	Tyr	Ser	Trp	
Ser	Gly 200	Ser	Ser	Pro	Leu	Thr 105	Asn	Phe	Leu	Asp	Pro 210	Ser	Asp	Ala	Phe	
Leu 215	Gly	Ala	Ser	Ser	Leu 220	Thr	Phe	Lys	Asn	Gly 225	Thr	Ala	Asn	Asp	Gly 230	
Leu	Val	Gly	Thr	Cys 235	Ser	sei) His	Leu	Gly 240	Met	Val	Ile	Arg	Asp 245	Asn	
Tyr	Arg	Met	Asn 250	His	Leu	Asp	Flu	Val 255	Asn	Gln	Val	Phe	Gly 260	Leu	Thr	
Ser	Leu	Phe 265	Glu	Thr	Ser	Pro	V 1 2 70	Ser	Val	Tyr	Arg	Gln 275	His	Ala	Asn	
Arg	Leu 280	Lys	Asn	Ala	Ser	Leu 285	\	١								
(2)	INF	ORMA'	rion	FOR	SEQ	ID I	10: 9	•								
	(i	() ()	QUENC A) LI B) T C) S' D) To	ENGTI YPE: IRANI	nuci nuci	047 l leic ESS:	ase acio unki	pa\i	rs \							
	(ii) MOI	LECU:	LE T	YPE:	DNA	(ger	nomi	=\							
	(ix	(2	ATURI A) Ni B) L	AME/			1016		\							
	(ix	(2	ATURI A) Ni B) L	AME/					\	\						
	(xi) SE	QUEN	CE D	ESCR	IPTI	ON:	SEQ :	ID N	o:√9	:					
GGA.	rccc	CGG '	TTCT	CCCG	GA A	GGAT	rcgg	G CG	ATGG	стсф	CAG	GACG	CGC (CCT	CGGCC	60
CAT	CAAC	CTG 1	AGAT	GAGA	AC A	M	rg Alet L	ys L	AG A. ys L	AG T ys S	r L	eu L	rc co eu P: 20	CC C	TC eu	110
GGC Gly	CTG Leu	GCC Ala -15	ATC Ile	GGT Gly	CTC Leu	GCC Ala	TCT Ser -10	CTC Leu	GCT Ala	GCC Ala	AGC Ser	CCT Pro -5	CTG Leu	ATC Ile	CAG Gln	158

GCC Ala	AGC Ser 1	ACC Thr	TAC\ Tyr	ACC	CAG Gln 5	ACC Thr	AAA Lys	TAC Tyr	CCC Pro	ATC Ile 10	GTG Val	CTG Leu	GCC Ala	CAC His	GGC Gly 15	206
ATG Met	CTC Leu	GGC Gly	TTC Phe	GAO Asp 20	AAC Asn	ATC Ile	CTC Leu	GGG Gly	GTC Val 25	GAC Asp	TAC Tyr	TGG Trp	TTC Phe	GGC Gly 30	ATT Ile	254
		GCC Ala														302
AGC Ser	CAG Gln	TTG Leu 50	GAC Asp	ACC Thr	TCG Ser	GAA 1u	GTC Val 55	CGC Arg	GGC Gly	GAG Glu	CAG Gln	TTG Leu 60	CTG Leu	CAA Gln	CAG Gln	350
GTG Val	GAG Glu 65	GAA Glu	ATC Ile	GTC Val	GCC Ala	CTO Leu 70	AGC Ser	GGC Gly	CAG Gln	CCC Pro	AAG Lys 75	GTC Val	AAC Asn	CTG Leu	ATC Ile	398
GGC Gly 80	CAC His	AGC Ser	CAC His	GGC Gly	GGG Gly 85	CCG Pro	A C	ATC Ile	CGC Arg	TAC Tyr 90	GTC Val	GCC Ala	GCC Ala	GTA Val	CGT Arg 95	446
CCC Pro	GAC Asp	CTG Leu	ATC Ile	GCT Ala 100	TCC Ser	GCC Ala	ACC Thr	AGC	GTC Val 105	GGC Gly	GCC Ala	CCG Pro	CAC His	AAG Lys 110	GGT Gly	494
		ACC Thr														542
		GTC Val 130														590
		TCC Ser														638
GAG Glu 160	TCG Ser	CTG Leu	AAC Asn	AGC Ser	GAG Glu 165	GGT Gly	GCC Ala	GCG Ala	CGC Arg	TTC Phe 170	AAC Asn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro 175	686
		ATC Ile														734
		AGC Ser														782
			Ser										Thr		AAG Lys	830
		Thr					Leu					se ∖ r			CTG Leu	878

GGC ATG GTG ATC CGC GAC AAC TAC CGG ATG AAC CAC CTG GAC GAG GTG 926 Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val AAC CAG GTC TTC GGC CTC ACC AGC CTG TTC GAG ACC AGC CCG GTC AGC 974 Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser GTC TAC CGC CAG CAG GCC AAC CGC CTG AAG AAC GCC AGC CTG Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu 1016 275 280 TAGGACCCCG GCCGGGCCT\CGGCCCGGGC C 1047 (2) INFORMATION FOR SEQ ID NO: 10: (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 311 amino acids
(B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protection (xi) SEQUENCE DESCRIPTION; SEQ ID NO: 10: Met Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -20 Ser Leu Ala Ala Ser Pro Leu Ile \$1n Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile to Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Val Set Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln Val Alu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile Gly Hit Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val\Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Ser Thr 140 145 Gly Thr Gln Asn Ser Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly

Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp 190 Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe 205 Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser\Ser His Leu Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr 255 Ser Leu Phe Glu Thr Ser Pro\Val Ser Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu (2) INFORMATION FOR SEQ ID NO: 11: (i) SEQUENCE CHARACTERISTICS (A) LENGTH: 1049 base palrs (B) TYPE: nucleic acid (C) STRANDEDNESS: unknown (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: DNA (genomic) (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 85..1017 (ix) FEATURE: (A) NAME/KEY: mat_peptide (B) LOCATION: 163. . 1017 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1/4: GGATCCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG CCAGGACGCG CCCCTCGGCC 60 CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG TCT CTG CTC CCC CTC Met Lys Lys Lys Set Leu Leu Pro Leu GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln ~15 GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC 207 Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Lau Ala His Gly ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT

Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile

20

			١													
			TTG Leu 35													303
AGC Ser	CAG Gln	TTG Leu 50	GAC Asp	ACC\ Thr	TCG	GAA Glu	GTC Val 55	CGC Arg	GGC Gly	GAG Glu	CAG Gln	TTG Leu 60	CTG Leu	CAA Gln	CAG Gln	351
GTG Val	GAG Glu 65	GAA Glu	ATC Ile	GTC Val	GQC Ald	CTC Leu 70	AGC Ser	GGC Gly	CAG Gln	CCC Pro	AAG Lys 75	GTC Val	AAC Asn	CTG Leu	ATC Ile	399
GGC Gly 80	CAC His	AGC Ser	CAC His	GGC Gly	GGG Gly 85	PCG PCG	ACC Thr	ATC Ile	CGC Arg	TAC Tyr 90	GTC Val	GCC Ala	GCC Ala	GTA Val	CGT Arg 95	447
CCC Pro	GAC Asp	CTG Leu	ATC Ile	GCT Ala 100	TCC Ser	GCC Ala	ACC	AGC Ser	GTC Val 105	GGC Gly	GCC Ala	CCG Pro	CAC His	AAG Lys 110	GGT Gly	495
TCG Ser	GAC Asp	ACC Thr	GCC Ala 115	GAC Asp	TTC Phe	CTG Leu	Ar	CAG Gln 120	ATC Ile	CCA Pro	CCG Pro	GGT Gly	TCG Ser 125	GCC Ala	GGC Gly	543
GAG Glu	GCA Ala	GTC Val 130	CTC Leu	TCC Ser	GGG Gly	CTG Leu	GTC Val 135	AAC Asn	AGC Ser	CTC Leu	GGC Gly	GCG Ala 140	CTG Leu	ATC Ile	AGC Ser	591
			AGC Ser													639
			AAC Asn													687
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr 180	TCG Ser	GCC Ala	TGC Cys	GGC Gly	GAA Glu 185	GGC	GCC Ala	TAC Tyr	AAG Lys	GTC Val 190	AAC Asn	735
GGC Gly	GTG Val	AGC Ser	TAT Tyr 195	TAC Tyr	TCC Ser	TGG Trp	AGC Ser	GGT Gly 200	TCC Ser	TCC Ser	CCG Pro	CTG Leu	ACC Thr 205	AAC Asn	TTC Phe	783
CTC Leu	GAT Asp	CCG Pro 210	AGC Ser	GAC Asp	GCC Ala	TTC Phe	CTC Leu 215	GGC Gly	GCC Ala	TCG Ser	rcc Ser	CTG Leu 220	ACC Thr	TTC Phe	AAG Lys	831
			GCC Ala													879
			ATC Ile								His					927
			TTC Phe												Ser	975

GTC TAC CGC CAG CAC CAC AGC CTG AAG AAC GCC AGC CTG Val Tyr Arg Gln His A 280

1017

TAGGACCCCG GCCGGGGCCT CGGCCCGGGC CC

1049

(2) INFORMATION FOR SEQ IN NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 311 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: lindar

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION; SEQ ID NO: 12: Met Lys Lys Lys Ser Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -20 Ser Leu Ala Ala Ser Pro Leu Il Sln Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Caly Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly Ger Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile Gly\His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg Pro App Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr Gly Thr Gln Asn Leu Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Lew Thr Phe Lys Asn Gly Thr Ala Asn Asp Gly

Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn 240 Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu 280 (2) INFORMATION FOR SEQ ID NO: 13: (i) SEQUENCE CHARACTERISTICS (A) LENGTH: 1050 base pairs (C) STRANDEDNESS: unknow (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: DNA (genomia) (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 85..1017 (ix) FEATURE: (A) NAME/KEY: mat peptide (B) LOCATION: 163..1017 (xi) SEQUENCE DESCRIPTION: SEO ID NO: \13: GGATCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC 60 CCATCAACCT GAGATGAGAA CAAC ATG AAG AAG AAG TCT CTG CTC CCC CTC

Met Lys Lys Ser Leu Leu Pro Leu -26 -25 GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG 159 Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln -15 GCC AGC ACC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Leu Ala His Gly ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile CCC AAC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GGC Pro Asn Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly 303 AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG CAG TTG CTG CAA CAG 351 Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln

		GAA Glu														399
		AGC Ser														447
		CTG Leu														495
		ACC Thr														543
GAG Glu	GCA Ala	GTC Val 130	CTC Leu	TCC Ser	GGG Gly	CTG Leu	GTC Val 135	AAC Asn	AGC Ser	CTC Leu	GGC Gly	GCG Ala 140	CTG Leu	ATC Ile	AGC Ser	591
TTC Phe	CTT Leu 145	TCC Ser	AGC Ser	GGC Gly	GGC Gly	ACC Thr 150	GG#	ACG Thr	CAG Gln	AAT Asn	TTA Leu 155	CTG Leu	GGC Gly	TCG Ser	CTG Leu	639
GAG Glu 160	TCG Ser	CTG Leu	AAC Asn	AGC Ser	GAG Glu 165	GGT Gly	GCC Ala	GOG Ala	CGC Arg	TTC Phe 170	AAC Asn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro 175	687
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr 180	TCG Ser	GCC Ala	TGC Cys	GGC Gly	GAA Glu 185	GGC Gly	GCT Ala	TAC Tyr	AAG Lys	GTC Val 190	AAC Asn	735
		AGC Ser														783
CTC Leu	GAT Asp	CCG Pro 210	AGC Ser	GAC Asp	GCC Ala	TTC Phe	CTC Leu 215	GGC Gly	GCC Ala	TCG	TCG Ser	CTG Leu 220	ACC Thr	TTC Phe	AAG Lys	831
		ACC Thr														879
		GTG Val									His					927
AAC Asn	CAG Gln	GTC Val	CTC Leu	GGC Gly 260	CTC Leu	ACC Thr	AGC Ser	CTG Leu	TTC Phe 265	GAG Glu	ACC Thr	AGC Ser	CCG Pro	GTC Val 270	Ser	975
		CGC Arg							Lys							1017
TAG	GACC	CCG	GCCG	GGGC	ст с	GGCC	CGGG	c cc	G		/					1050

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 311 amino acids
 - (B) TYPE amino acid
- (D) TOPONOGY: linear
- (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Met Lys Lys Ser Deu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala -26 -25

Ser Leu Ala Ala Ser Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr

Lys Tyr Pro Ile Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile 15

Leu Gly Val Asp Tyr Trp Phe Gly Ile Pro Asn Ala Leu Arg Arg Asp 25 30 35

Gly Ala Gln Val Tyr Val Tha Glu Gly Ser Gln Leu Asp Thr Ser Glu

Val Arg Gly Glu Gln Leu Leu din Gln Val Glu Glu Ile Val Ala Leu

Ser Gly Gln Pro Lys Val Asn Lev Ile Gly His Ser His Gly Gly Pro

Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala

Thr Ser Val Gly Ala Pro His Lys Gly Ser Asp Thr Ala Asp Phe Leu

Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu

Val Asn Ser Leu Gly Ala Leu Ile Ser Ahe Leu Ser Ser Gly Gly Thr 135

Gly Thr Gln Asn Leu Leu Gly Ser Leu Gl ψ Ser Leu Asn Ser Glu Gly

Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln\Gly Ile Pro Thr Ser Ala 175

Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp

Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe

Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly\Thr Ala Asn Asp Gly

Leu Val Gly Thr Cys Ser Ser His Leu Gly Met $\sqrt[4]{al}$ Ile Arg Asp Asn 235 240 245

Tyr Arg Met Ast His Leu Asp Glu Val Asn Gln Val Leu Gly Leu Thr 250 255 Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Gln His Ala Asn 265 270 Arg Leu Lys Asn Ala Ser Leu 280 285 (2) INFORMATION FOR SEQ ID NO: 15: (i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 1049 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: unknown (D) TOPOLOGY: unknown (ii) MOLECULE TYPE: DNA (genomic) (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 85..1017 (ix) FEATURE: (A) NAME/KEY: mat pertide (B) LOCATION: 163...1017 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15: GGATCCCCG GTTCTCCCGG AAGGATTCGG GCGATGGCTG GCAGGACGCG CCCCTCGGCC CCATCAACCT GAGATGAGAA CAAC ATG AAG \AAG AAG TCT CTG CTC CCC CTC
Met Lys Lys Ser Leu Leu Pro Leu 111 -26 -25 -20 GGC CTG GCC ATC GGT CTC GCC TCT CTC GCT GCC AGC CCT CTG ATC CAG Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Ser Pro Leu Ile Gln -15 GCC AGC AGC TAC ACC CAG ACC AAA TAC CCC ATC GTG CTG GCC CAC GGC Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro the Val Leu Ala His Gly ATG CTC GGC TTC GAC AAC ATC CTT GGG GTC GAC TAC TGG TTC GGC ATT 255 Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp Phe Gly Ile CCC AGC GCC TTG CGC CGT GAC GGT GCC CAG GTC TAC GTC ACC GAA GGC 303 Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly AGC CAG TTG GAC ACC TCG GAA GTC CGC GGC GAG dAG TTG CTG CAA CAG Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu Gln Gln 351 GTG GAG GAA ATC GTC GCC CTC AGC GGC CAG CCC AAG GTC AAC CTG ATC Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lya Val Asn Leu Ile 65

GGC Gly 80	CAC His	AGC Ser	CAC Hi	GGC Gly	GGG Gly 85	CCG Pro	ACC Thr	ATC Ile	CGC Arg	TAC Tyr 90	GTC Val	GCC Ala	GCC Ala	GTA Val	CGT Arg 95	447	
CCC Pro	GAC Asp	CTG Leu	ATC Ile	ACT AA 100	TCC Ser	GCC Ala	ACC Thr	AGC Ser	GTC Val 105	GGC Gly	GCC Ala	CCG Pro	CAC His	AGG Arg 110	GGT Gly	495	
TCG Ser	GAC Asp	ACC Thr	GCC Ala 115	GAC Asp	TTC	CTG Leu	CGC Arg	CAG Gln 120	ATC Ile	CCA Pro	CCG Pro	GGT Gly	TCG Ser 125	GCC Ala	GGC Gly	543	
GAG Glu	GCA Ala	GTC Val 130	CTC Leu	TCC Ser	G17 GG	CTG Leu	GTC Val 135	AAC Asn	AGC Ser	CTC Leu	GGC Gly	GCG Ala 140	CTG Leu	ATC Ile	AGC Ser	591	
TTC Phe	CTT Leu 145	TCC Ser	AGC Ser	GGC Gly	GGC Gly	ACC Thir 150/	GGT Gly	ACG Thr	CAG Gln	AAT Asn	TTA Leu 155	CTG Leu	GGC Gly	TCG Ser	CTG Leu	639	
GAG Glu 160	TCG Ser	CTG Leu	AAC Asn	AGT Ser	GAG Glu 165	GGT Gly	CCC Ala	GCG Ala	CGC Arg	TTC Phe 170	AAC Asn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro 175	687	
CAG Gln	GGC Gly	ATC Ile	CCC Pro	ACC Thr 180	TCG Ser	GCC Ala	TGC Cys	GGC Gly	GAA Glu 185	GGC Gly	GCT Ala	TAC Tyr	AAG Lys	GTC Val 190	AAC Asn	735	
				TAC Tyr												783	
CTC Leu	GAT Asp	CCG Pro 210	AGC Ser	GAC Asp	GCC Ala	TTC Phe	CTC Leu 215	GGC Gly	GCC	TCG Ser	TCG Ser	CTG Leu 220	ACC Thr	TTC Phe	AAG Lys	831	
AAC Asn	GGC Gly 225	ACC Thr	GCC Ala	AAC Asn	GAC Asp	GGC Gly 230	CTG Leu	GTC Val	G1 X	ACC Thr	TGC Cys 235	AGT Ser	TCG Ser	CAC His	CTG Leu	879	
				CGC Arg												927	
AAC Asn	CAG Gln	GTC Val	CTC Leu	GGC Gly 260	CTC Leu	ACC Thr	AGC Ser	CTG Leu	TTC Phe 265	GAG Glu	ACC Thr	AGC Ser	CCG Pro	GTC Val 270	AGC Ser	975	
GTC Val	TAC Tyr	CGC Arg	CAG Gln 275	CAC His	GCC Ala	AAC Asn	CGC Arg	CTG Leu 280	AAG Lys	AAC Asn	GCC A a	AGC Ser	CTG Leu 285			1017	
TAG	GACC	CCG (GCCG	GGGC	CT C	GCC	CGGG	c cc			\					1049	

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 311 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met Lys Lys Lys der Leu Leu Pro Leu Gly Leu Ala Ile Gly Leu Ala Ser Leu Ala Ala Sex Pro Leu Ile Gln Ala Ser Thr Tyr Thr Gln Thr Lys Tyr Pro Ile Val Deu Ala His Gly Met Leu Gly Phe Asp Asn Ile Leu Gly Val Asp Tyr Trp\Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp
25 30 35 Gly Ala Gln Val Tyr Val Thr Glu Gly Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gln Leu Leu 11 Gln Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asd Lev Ile Gly His Ser His Gly Gly Pro 75 80 85 Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro His Arg Gly Ser Asp Thr Ala Asp Phe Leu Arg Gln Ile Pro Pro Gly Ser Ala Gly Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu Ile Ser Phe Leu Ser Ser Gly Gly Thr 135 145 Gly Thr Gln Asn Leu Leu Gly Ser Leu Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr Pro Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Asn Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys Asn Gly thr Ala Asn Asp Gly Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val Asn Gln Val Leu Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr Arg Cln His Ala Asn Arg Leu Lys Asn Ala Ser Leu

(2)	INFO	RMAT	три	FOR	SEQ	ID N	10: 1	7:								
	(i)	(E	UENC) LE) TY :) ST) TO	NGTH RE: RAND	: 10 nucl EDNE	49 b eic SS:	ase acid unkn	pair I	s							
	(ii)	MOL	ECUL	Е Т	PE:	DNA	(gen	omic	:)							
	(ix)		TURE NA L) LC	ME/K			.017									
	(ix)	FEA (A	TURE) NA) LC	ME/K	EY:	mat_ 63	pept	ide								
	(xi)	SEÇ	UENC	E DE	SCRI	PAIC	N: S	EQ I	D NO): 17	· :					
GGAI	cccc	CG G	TTCT	ccce	G AF	reey	TGGG	GCG	ATGO	CTG	GCAG	GACG	GCG C	CCCI	CGGCC	60
CCAT	CAAC	CT G	AGAT	'GAGA	A CA	1	26 I	ys I	AAG A	AAG T ∴ys S	CT C Ser I	eu I	TC C Leu I -20	ccc c	TC eu	111
GGC Gly	CTG Leu	GCC Ala -15	ATC Ile	GGT Gly	CTC Leu	GCC Ala	TCT Ser -10	CTC Leu	GCT Ala	GCC Ala	AGC Ser	CCT Pro -5	CTG Leu	ATC Ile	CAG Gln	159
	AGC Ser 1															207
ATG Met	CTC Leu	GGC Gly	TTC Phe	GAC Asp 20	AAC Asn	ATC Ile	CTT Leu	GGG/	GTC Val 25	GAC Asp	TAC Tyr	TGG Trp	TTC Phe	GGC Gly 30	ATT Ile	255
	AGC Ser															303
AGC Ser	CAG Gln	TTG Leu 50	GAC Asp	ACC Thr	TCG Ser	GAA Glu	GTC Val 55	CGC Arg	GGC Gly	GAG Glu	CAG Gln	TTG Leu 60	CTG Leu	CAA Gln	CAG Gln	351
	GAG Glu 65															399
	CAC His										\Val					447
CCC Pro	GAC Asp	CTG Leu	ATC Ile	GCT Ala 100	TCC Ser	GCC Ala	ACC Thr	AGC Ser	GTC Val 105	GGC Gly	dcc Ala	CCG Pro	CAC His	AAG Lys 110	GGT Gly	495

			GCC Ala 115													543
			CTC Leu												AGC Ser	591
TTC Phe	CTT Leu 145	TCC Ser	AGC Ser	GGC\ Gly	GGC Gly	ATC Ile 150	GGT Gly	ACG Thr	CAG Gln	AAT Asn	TTT Phe 155	CTG Leu	GGC Gly	TCG Ser	CTG Leu	639
GAG Glu 160	TCG Ser	CTG Leu	AAC Asn	AGC Ser	GAG G1 165	GGT Gly	GCC Ala	GCG Ala	CGC Arg	TTC Phe 170	AAC Asn	GCC Ala	AAG Lys	TAC Tyr	CCG Pro 175	687
Gln	Gly	Ile	CCC Pro	Thr 180	Ser	Ata	Cys	Gly	Glu 185	Gly	Ala	Tyr	Lys	Val 190	Asn	735
GGC Gly	GTG Val	AGC Ser	TAT Tyr 195	TAC Tyr	TCC Ser	TGG	AGC Ser	GGT Gly 200	TCC Ser	TCG Ser	CCG Pro	CTG Leu	ACC Thr 205	AAC Asn	TTC Phe	783
			AGC Ser													831
			GCC Ala													879
GGC Gly 240	ATG Met	GTG Val	ATC Ile	CGC Arg	GAC Asp 245	AAC Asn	TAC Tyr	CGG\ Arg	ATG Met	AAC Asn 250	CAC His	CTG Leu	GAC Asp	GAG Glu	GTG Val 255	927
AAC Asn	CAG Gln	GTC Val	TTC Phe	GGC Gly 260	CTC Leu	ACC Thr	AGC Ser	CTG Leu	TTC Phe 265	GAG Glu	ACC Thr	AGC Ser	CCG Pro	GTC Val 270	AGC Ser	975
			CAG Gln 275													1017
TAG	GACC	CCG	GCCG	GGGC	CT C	GGCC	CGGG	c cc		/	\					1049
(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	18:								
		(SEQU A) L B) T D) T	ENGT YPE:	H: 3	11 a no a	mino cid					\				
			LECU						ID N	0: 1	8:					
	Lys -25		Lys	Ser	Leu	Leu -20		Leu	Gly	Leu	Ala -15		Gly	Leu	Ala	
Ser -10		Ala	Ala	Ser	Pro		Ile	Gln	Ala	Ser 1		Tyr	Thr	Gln 5	Thr	

Lys Tyr Pro I $\frac{1}{10}$ e Val Leu Ala His Gly Met Leu Gly Phe Asp Asn Ile 15 Leu Gly Val Asp\Tyr Trp Phe Gly Ile Pro Ser Ala Leu Arg Arg Asp Gly Ala Gln Val Tyr Val Thr Glu Gly Ser Gln Leu Asp Thr Ser Glu Val Arg Gly Glu Gla Leu Leu Gln Gln Val Glu Glu Ile Val Ala Leu Ser Gly Gln Pro Lys Val Asn Leu Ile Gly His Ser His Gly Gly Pro Thr Ile Arg Tyr Val Ala Ala Val Arg Pro Asp Leu Ile Ala Ser Ala Thr Ser Val Gly Ala Pro $H \searrow S$ Lys Gly Ser Asp Thr Ala Asp Phe Leu 105 Arg Gln Ile Pro Pro Gly Sel Ala Gly Glu Ala Val Leu Ser Gly Leu Val Asn Ser Leu Gly Ala Leu The Ser Phe Leu Ser Ser Gly Gly Ile Gly Thr Gln Asn Phe Leu Gly Set Leu Glu Ser Leu Asn Ser Glu Gly Ala Ala Arg Phe Asn Ala Lys Tyr to Gln Gly Ile Pro Thr Ser Ala Cys Gly Glu Gly Ala Tyr Lys Val Ash Gly Val Ser Tyr Tyr Ser Trp Ser Gly Ser Ser Pro Leu Thr Asn Phe Leu Asp Pro Ser Asp Ala Phe Leu Gly Ala Ser Ser Leu Thr Phe Lys Ash Gly Thr Ala Asn Asp Gly 215 Leu Val Gly Thr Cys Ser Ser His Leu Gly Met Val Ile Arg Asp Asn Tyr Arg Met Asn His Leu Asp Glu Val Asn Glm Val Phe Gly Leu Thr Ser Leu Phe Glu Thr Ser Pro Val Ser Val Tyr\Arg Gln His Ala Asn Arg Leu Lys Asn Ala Ser Leu

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs (B) TYPE: nucleic acid

(B) TYPE: nucleic acid(C) STRANDEDNESS: unknown

30

25

27

(D) TOPOLOGY: linear

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
GCGC	AATTAA CCCTCACTAA AGGGAACAAA
(2)	INFORMATION FOR SEA ID NO: 20:
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 25 base pairs (B) TYPE: nucley a acid (C) STRANDEDNESS) unknown (D) TOPOLOGY: 1 mear
	(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: \(\text{desc} = \text{"synthetic DNA"} \)
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:
	CGCAGA ATNNNCTGGG CTCGC INFORMATION FOR SEQ ID NO: 21/2:
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 27 base pairs (B) TYPE: nucleic acid (C) STRANDEDDESS: unknown (D) TOPOLOGY: linear
	(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = synthetic DNA"
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
GCGT/	AATACG ACTCACTATA GGGCGAA

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"